

Isabel Zavian

izavian@berkeley.edu | [Website](#) | (818) 517-4250 | [GitHub](#) | [LinkedIn](#)

EDUCATION

University of California, Berkeley

Aug 2018 - May 2022

B.A. Economics & B.A. Data Science w/ Specialization in Business and Industrial Analytics

Certificate of Entrepreneurship and Technology

Relevant Coursework: Micro-/Macroeconomics, Econometrics, Data Structures, Structure & Interpretation of Computer Programs, Linear Algebra & Differential Equations, Financial Economics, Psychology & Economics, Probability & Machine Learning for Data Science, Applied Data Science w/ Venture Applications, Advanced Business Analytics, Corporate Financial Statement Analysis

EXPERIENCE

PlanetBids, Inc. | *Data Scientist*

May 2022 – Present

- Analyzed and cleaned unstructured Big Data using SQL and Python libraries to improve collection process by 32%
- Developing text mining pipelines and using toolkits, such as NLTK, for complex text classification with NLP on real time data
- Building a cost estimation model with statistical ML models, such as Naïve Bayes, & SVMs, for government bidding projects

Honda R&D Americas, LLC | *Data Analyst Intern*

Jun 2021 – Aug 2021

- Reduced Big Data collection time from 5 different sources by 2 hours through building pipelines from S3 buckets via PySpark
- Optimized data visualization procedures by developing summary dashboards, saving 12+ hours a week of complex data analysis
- Created enrichment modules to filter and balance unstructured datasets containing over 75 prototyped model features
- Built machine learning models (LogReg, Random Forests, etc.) and improved accuracy by 30% through statistical methods

Colt Technology Services | *Product Manager Intern*

Jun 2021 – Aug 2021

- Performed deep market and competition research to deliver an improved AI customer experience on network performance
- Improved network efficiency by 23% by developing an algorithm that automatically suggests an optimal bandwidth based on various QOS levels and by building a web application with metric visualizations via Anvil Works
- Remotely managed and led a large team in five different time zones from product conception through launch

Honda R&D Americas, LLC | *Machine Learning Consultant*

Jan 2021 – May 2021

- Built a variety of models and algorithms (i.e. Markov, KNNs) to predict a vehicle's route, next location, and stop duration
- Cloud processed, stored, and trained millions of records of data to develop a destination and dwell time simulation model that provides potential on-demand services and vehicle repairs
- Increased the model's interpretability and accuracy to 86% by using Python visualization libraries, DBScan, and XGBoost

LEADERSHIP

Women in Computing and Data Science at Berkeley | *Founder and President*

Jan 2021 – May 2022

- Established the club's officer team, overall structure, semester agenda, and Machine Learning courses for committees
- Advised company projects in Python to conduct deep data analysis via cleaning, visualizations, etc., & drive business decisions
- Created partnerships with university departments (DS, CS, etc.) and professional organizations to provide mentorship and career opportunities for members and have access to worldwide hackathons and conferences

Undergraduate Economics Association | *President*

Dec 2020 – May 2021

- Led and provided direction to the club by establishing objectives and delegating tasks for the officer team and 300+ members
- Continuously monitored the club's progress towards its goals and trained officers to carry out their responsibilities

PROJECTS

U.S. Asthma Rates

Nov 2021 – Dec 2021

- Analyzed causal effects of PM2.5 concentrations on U.S. asthma prevalence via causal inference, multiple hypothesis testing, etc

Spam/Ham

Nov 2020 – Dec 2020

- Created a classifier with 96% accuracy to distinguish spam emails by minimizing overfitting and validating model performance, using sklearn libraries to fit models, analyzing precision-recall curves, and utilizing text data with feature engineering

Food Safety

Sep 2020 – Oct 2020

- Used Python, SQL, and Pandas to work with data at different levels of granularity, explore characteristics and distributions of individual variables, and identify the type of data collected, missing values, anomalies, etc., to analyze food safety in SF

SKILLS

Skills: Python, Java, PySpark, R, SQL, CSS, Javascript, HTML, Apache Superset, Microsoft Office, Written & Verbal Communication, Time Management, Data Analysis & Visualization, Tableau, Machine Learning, A/B Testing, Statistics

Libraries: Pandas, NumPy, Matplotlib, Seaborn, SciPy, SciKit-Learn, TensorFlow, Imb-Learn, XGBoost, PyTorch